REMARKS

Claims 1-4, 7-19, and 21-38 are pending in the application. Reconsideration of the above-identified application in view of the following remarks is respectfully requested.

Claim Rejections – 35 USC 103

In this section of the Office Action, claims 1-4, 7-19, and 21-38 were rejected under 35 USC 103(a) as being as being unpatentable by Ferrel et al., US006199082B1 (Ferrell '082) in view of Fitzsimons et al., US006708189B1 (Fitzsimons '189), further in view of Wyler US20050055420A1 (Wyler '420).

Favorable reconsideration of this rejection is respectfully requested since, as will be shown below, the above-amended claims are both novel and inventive over the prior art cited by the Examiner.

The present invention relates to a system which automatically carries out intelligent structure analysis of document structure in order to understand the separate parts of a document and the internal structures of those parts. The system is then used to parse documents, using a marking system to mark the structure on the documents and then using the marking to republish the original document in a second format. This is discussed in the present application, specifically at the first paragraph of page 8, wherein it is noted that "the present invention does not merely convert the data from one format to a different format, but instead is able to also provide at least a basic structure for organizing the data into the Web page."

The system of the present invention is detailed further in the specification, at page 11, third paragraph – page 17, wherein it is stated that, according to the

present invention, data which is in a <u>digital format can be converted to a basic internal format</u>, which is preferably XML, although substantially any other type of mark-up language could also be used. Next, <u>intelligent structure analysis is performed, in order to recognize and define the structures and objects</u> contained in the newspaper data, particularly with regard to each page of the newspaper. More preferably, newspaper objects are classified as one of a plurality of specific entities, which optionally and most preferably are selected from the group consisting of an article, an advertisement, a picture which is not otherwise associated with an article or advertisement, and general data, which covers information that does not fit into one of the preceding categories. Recognition of the hierarchy for newspaper objects is an important part of the process. The process of intelligent structure analysis enables the newspaper data to be converted to a series of objects, for more efficient search and retrieval through the Internet or other network. After the process of intelligent structure analysis has been completed, <u>encoding of the object data is performed</u>. This results in a set of enhanced, structured files which combine the original image of the data with the text and encoding information. <u>Each such file thus preferably maintains the visual aspects of the newspaper layout, while enabling far greater functionality to be available through the Web page version of the newspaper</u>. The internal format data is published in a plurality of different final formats by a publication server.

According to the present invention, typically, a document in an original, existing format is divided into pages, articles, pictures, pictures with articles, articles with headlines, articles with headlines and a by-line, articles with headlines, by-line, a picture, and the like. The document could be a newspaper, a company financial report, or any kind of structured document. The final published format is an interactive document, typically for use as a Web page. The automatic parsing of the

present embodiments allows the structural subunits, the blocks, of the original document to be the interactive objects of the Web page. Thus a particular block, comprising a headline, by-line, text and a photograph, is identified automatically as belonging together. The block is defined as a single object, and the text, etc., as internal structure thereof. The parser classifies the headline using an XML label, the by-line likewise, the text likewise, and the picture likewise, and then the publisher knows that the entire object should be presented visually as it appears in the original newspaper page. However it also knows how to set up the headline as a hypertext link to lead to a larger size version of the full text so that the text can be read. Likewise it knows how to set up the picture as a link that can be clicked on to lead to a larger view of the picture. The by-line may be treated separately to provide a link to other articles by the same author. The result is a fully interactive version of the original document.

In summary, the interactive version of a multi-element document is automatically arrived at by reverse engineering (parsing) of the document in the original, existing format, leading to XML labeling of the document parts to represent blocks and internal structure of the blocks, the labeling used as directions by a publisher to publish an interactive version of the document.

Ferrel '082 discloses a multimedia publishing system for <u>creating online publications,</u> wherein the contents and design aspects of the online publishing process are separated into distinct components before being transmitted to the customers' computers. Separation of the contents and the design aspects of the online publications facilitates efficient distribution of published documents by enabling the transmission of, for example, high-quality titles over low-speed communications links subject to loss of connectivity. As discussed in Ferrel, specifically at column 5, line 9,

since, for a particular online publication, the design of many titles remains static while the content changes for each new publication issue, layout designs for that online publication may be cached on a user's computer, thus minimizing the transmission of redundant data and optimizing bandwidth use. Thus, the system enables the content of each issue to be transmitted relatively quickly, since it contains only tagged components, not the actual pages and controls themselves. While the system to Ferrel enables high-quality publications to be sent to customer's computers (column 62, line 53), there is neither shown nor suggested a system whereby data in the form of a newspaper/having a plurality of pages in an original, existing format is converted into a final publication format, as in the present invention.

The Examiner's suggestion that Ferrel teaches automatically publishing data in the form of a newspaper has been noted. It should, however, be pointed out that Ferrel does not relate to publishing an on-line publication from data in the form of a newspaper having an original, existing format, as in the present invention. As discussed in Ferrel, specifically at column 7, line 60 – column 8, line 14, the publishing system disclosed receives data on the Microsoft Network, wherein the data is separated into content and design components. While the design component allows the content component to appear as a newspaper on a computer monitor and the content makes up the designed screens, no actual newspaper having an original, existing format, is employed.

Fitzsimons '189 discloses a computer file transfer system wherein a data file is received by an automated publishing system from a source, and the <u>data file is converted for use within a presentation space of the automated publishing system</u>. As discussed at column 12, line 60 – column 13, line 55 of the patent to Fitzsimons, the system converts a file, such as an MS Word document file or a digital image TIFF

file, created under a first format, into a corresponding data file under a second format. This enables communication between computers which must exchange files having different data structures. There is, however, neither shown nor suggested converting data in the form of a newspaper/having a plurality of pages in an original, existing format into a final publication format, as in the present invention.

Wyler '420 discloses a method and system for processing information received by a wireless device over a computer network, wherein the information received is converted into a form suitable for displaying on the wireless device. As discussed at sections [0455] – [0456] of the publication to Wyler, information is received over a computer network and at least some of the information is "parsed," in the sense that irrelevant information is removed. It should be noted that the use of the term "parsing" by Wyler to mean removing unwanted information differs significantly from the use of the term "parsing" in the present invention to mean reverse engineering of the original document, leading to XML labeling of the document parts to represent blocks and internal structure of the blocks, as discussed above. Finally, according to Wyler, at least some of the results are utilized to provide information suitable for display to a user on a wireless device. Wyler thus enables the information contained in a Web site to be presented on the display of a wireless device. The publication to Wyler is not concerned with the converting of data in the form of a newspaper/having a plurality of pages in an original, existing format into a final publication format, as in the present invention.

The Examiner has drawn our attention to the text of Wyler '420, specifically at page 14, paragraph [0360], through page 15, paragraph [0434], wherein there is a general description of format block (item 120) to include a means of automatic detection for an appropriate style. It should be noted that this format block is to be

utilizled by a user in order to choose a style in which to view a processed Webpage on a wireless device. As noted above, the system taught by Wyler '420 scans a Webpage and its information is transformed into objects, such as discussed at page 7, paragraphs [0157] – [0158]. The Webpage is parsed, as discussed at paragraphs [0164] – [0165], to remove irrelevant information, such as advertising banners and unrelated links, and the remaining information is transmitted for display on the wireless device. Wyler neither shows nor suggests, however, analyzing data in an original, existing format including a plurality of pages, as in the above-amended claims.

With reference to the text of Wyler '420, specifically at paragraphs [0455] – [456], information is received over a computer network and is processed so as to render the information in a form suitable for display on a wireless device. Figure 18a of this reference shows that information from a Website, such as that including banners, links, advertising, etc., is converted into a form suitable display on a wireless device, wherein only the news content is displayed. Similarly, in the text at paragraph [460] with regard to Figure 18b and in the text at paragraph [464] with regard to Figure 18c, it is noted that Web site information is processed such that only news content is displayed on a wireless device.

None of the cited documents either shows or suggests automatically analyzing an original, existing publication, such as a newspaper or a microfiche carried document, in such a reverse engineering process, in order to understand the separate parts of a document and the internal structures of those parts, as taught by the present invention. Additionally, as discussed in the present application, specifically at the last complete paragraph on page 10, "One of the advantages of the present invention is that archived data can optionally be easily integrated with current sources

of data, such that the newspaper which is published in the Web page format can represent a combination of current and previously published information. Second source 14 is shown as microfilm data, with an optional but preferred associated microfilm publisher 16 according to the present invention, for converting the microfilm data into a digital format, by converting the microfilm data to digital images." The cited references, either alone or in combination, are not able to provide such a system.

In contrast to the systems of the cited art, amended **Claim 1** recites "A method for automatically publishing data in a final publication format, wherein the data is in the form of a newspaper having an original, existing format including a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having an internal structure, the method comprising: automatically analyzing the data to decompose the predetermined layout of each page of the newspaper in the original, existing format into said plurality of blocks, each block representing an object, converting each object to an internal publication format, the format identifying said internal structure, and rendering said internal publication format to incorporate said objects and respective internal structures in the final publication format."

As explained above, and defined by amended claim 1, the present invention introduces the novel and inventive idea of a method where a publication in an original, existing format, such as a newspaper is automatically analyzed, in order to understand the separate parts of the published document and the internal structures of those parts, as described in 24, page 11: "Next, XML distiller module 18 preferably performs intelligent structure analysis, in order to be able to recognize and define the structures and objects contained in the newspaper data, particularly with regard to

each page of the newspaper. Examples of such structures and objects include, but are not limited to, articles, advertisements, titles, and so forth."

In contrast, Ferrel teaches a publication system wherein the publication process, carried out for the initial generation of the published document, is divided into a content aspect and a design aspect. However, Ferrel never suggests or even hints at the idea of automatically analyzing the data to decompose the layout of each page of an already generated newspaper in an original, existing format into the plurality of blocks, each block representing an object, converting each object to an internal publication format, the format identifying the internal structure, in such a reverse engineering process as taught by the present invention. Thus Ferrel deals with a totally different problem than the present invention and teaches a different solution than the present application. Fitzsimons also falls short of teaching or event hinting at such a method for automatically publishing data in a final publication format, wherein the data is in the form of a newspaper having an original, existing format including a plurality of pages, as taught by the present invention. As to Wyler, he discloses a system for processing Web site information including removing irrelevant information, so as to render the information in a form suitable for display on a wireless device, which teaches away from the present invention of automatically publishing data in a final publication format, wherein the data is in the form of a newspaper having an original, existing format including a plurality of pages.

It is thus respectfully believed that amended claim 1 is both novel and inventive over the prior art, and maintained that amended claim 1 is allowable.

Amended **Claim 19** recites "A system for automatically publishing newspaper data in a computerized format, the system comprising: (a) at least one source of newspaper data in an original, existing digital format, the newspaper format

comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having an internal structure; (b) a mark-up language distiller module for converting the data from said original, existing digital format to a mark-up language format, wherein said mark-up language distiller module automatically analyzes the newspaper data in said original, existing digital format to decompose the newspaper data into said plurality of blocks, each block with said internal structure representing an independent data object, each object having content and at least one attribute of the data, such that each object is converted to said mark-up language format; and (c) a publisher server for converting the data from said mark-up language format to a final publication format, said final publication format incorporating said blocks with the structure as objects."

As explained above, and defined by amended claim 19, the present invention introduces the novel and inventive idea of a system where a publication in an original, existing format such as a newspaper is automatically analyzed, by a mark-up language distiller module, in order to understand the separate parts of the published document and the internal structures of those parts.

Ferrel '082, as described above, teaches a publication system wherein the publication process, carried out for the initial generation of the published document, is divided into a content aspect and a design aspect. However, Ferrel never suggests or even hints at the idea of automatically analyzing the data in an original, existing digital format to decompose the layout of each page of the newspaper into the plurality of blocks, each block representing an object, converting each object to an internal publication format, the format identifying the internal structure, as taught by the present invention. Fitzsimons '189 also falls short of teaching or event hinting at such a system for automatically analyzing data in an original, existing digital format

and publishing data in a final publication format, as taught by the present invention. Similarly, Wyler teaches away from the present invention, by disclosing a system for processing Web site information including removing irrelevant information, so as to render the information in a form suitable for display on a wireless device.

It is thus respectfully believed that amended claim 19 is both novel and inventive over the prior art, and maintained that amended claim 19 is allowable.

Amended **Claim 23** recites "A method for automatically publishing data in a final publication format, the data having an original, existing format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having an internal structure the method comprising: automatically analyzing the data to decompose the data in the original, existing format into a plurality of objects, each object corresponding to one of said blocks; preparing a list of text and/or graphic elements for each object; determining properties of each element, including determining visibility and overlap characteristics for each graphic element within said object; recognizing structural layout properties of the data in an original format; converting each object to an internal publication format; and rendering said internal publication format in the final publication format, said final publication format presenting said blocks as said independently standing objects incorporating said internal structure."

As explained above, and defined by amended claim 23, the present invention introduces the novel and inventive idea of a method for automatically publishing where an original, existing multiple paged document having an original, existing format, is automatically analyzed, in order to understand the separate blocks of the document, thus decomposing the data in the original, existing format into a plurality of objects, each object corresponding to one of the blocks.

Ferrel '082, as described above, teaches a publication system wherein the publication process, carried out for the initial generation of the published document, is divided into a content aspect and a design aspect. However, Ferrel never suggests or even hints at the idea of automatically analyzing the data to decompose the layout of each page of the document in an original, existing format into the plurality of blocks, as taught by the present invention. Fitzsimons also falls short of teaching or event hinting at such a system for automatically publishing data in a final publication format, as taught by the present invention. As to Wyler, his system for processing Web site information including removing irrelevant information, so as to render the information in a form suitable for display on a wireless device, teaches away from the present invention of automatically publishing data in a final publication format, wherein the data is in the form of a newspaper having an original, existing format including a plurality of pages.

It is thus respectfully believed that amended claim 23 is both novel and inventive over the prior art, and maintained that amended claim 23 is allowable.

Amended **Claim 24** recites "A method for automatically publishing data in a final publication format, the data having an original, existing format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, the method comprising: automatically analyzing the data in the original, existing format to decompose the data into a plurality of objects, said objects corresponding to said blocks; preparing a list of text and/or graphic elements for each object; determining properties of each element, including determining a special characteristic for each text element; recognizing structural layout properties of the data in an original format; converting each object to an

internal publication format; and rendering said internal publication format in the final publication format such as to include said recognized structure in said objects."

As explained above, and defined by amended claim 24, the present invention introduces the novel and inventive idea of a method for automatically publishing where a multiple paged data in an original, existing format, is automatically analyzed, in order to understand the separate blocks of the data, thus decomposing the data in the original format into a plurality of objects, each object corresponding to one of the blocks.

Ferrel '082, as described above, teaches a publication system wherein the publication process, carried out for the initial generation of the published document, is divided into a content aspect and a design aspect. However, Ferrel never suggests or even hints at the idea of automatically analyzing the data in the original, existing format to decompose the layout of each page of the data into the plurality of blocks, as taught by the present invention. Fitzsimons also falls short of teaching or event hinting at such a system for automatically publishing data in a final publication format, , wherein the data has an original, existing format including a plurality of pages, as taught by the present invention. As to Wyler, he discloses a system for processing Web site information including removing irrelevant information, so as to render the information in a form suitable for display on a wireless device, which teaches away from the present invention of automatically publishing data in a final publication format, wherein the data is in an original, existing format including a plurality of pages.

It is thus respectfully believed that amended claim 24 is both novel and inventive over the prior art, and maintained that amended claim 24 is allowable.

Amended **Claim 25** recites "A method for automatically publishing data in a final publication format, wherein the data is in the form of a newspaper, the newspaper having an original, existing format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having structural layout properties, the method comprising: automatically analyzing the data to decompose the data in the original, existing format into a plurality of objects, said objects corresponding to said independently standing blocks; preparing a list of text and/or graphic elements for each object; determining properties of each element; recognizing said structural layout properties of the data in an original format; determining each text segment for each object; building a text block from a plurality of aligned text segments; converting each object to an internal publication format; and rendering said internal publication format in the final publication format to comprise said blocks as objects incorporating said structural layout properties."

As explained above, and defined by amended claim 25, the present invention introduces the novel and inventive idea of a method wherein a publication in an original, existing format, such as a newspaper is automatically analyzed, in order to understand the separate parts of the original, existing document and the internal structures of those parts.

Ferrel '082, as described above, teaches a publication system wherein the publication process, carried out for the initial generation of the published document, is divided into a content aspect and a design aspect. However, Ferrel never suggests or even hints at the idea of automatically analyzing a newspaper, to decompose the layout of each page of the newspaper in the original, existing format into the plurality of blocks, each block representing an object, converting each object to an internal

publication format, the format identifying the internal structure, as taught by the present invention. Fitzsimons '189 also falls short of teaching or event hinting at such a method for automatically publishing data, the data in the form of a newspaper in an original, existing format, in a final publication format, as taught by the present invention. As to Wyler, he discloses a system for processing Web site information including removing irrelevant information, so as to render the information in a form suitable for display on a wireless device, which teaches away from the present invention of automatically publishing data in a final publication format, wherein the data is in the form of a newspaper having an original, existing format including a plurality of pages.

It is thus respectfully believed that amended claim 25 is both novel and inventive over the prior art, and maintained that amended claim 25 is allowable.

In light of the above, it is submitted that above-amended claims 1, 19, 23, 24, and 25 are patentable over Ferrel '082 in view of Fitzsimons '189, further in view of Wyler '420, and are, therefore, allowable in accordance with USC 103(a).

Claims (2-4, 7-18, and 26-38) and (21-22), are believed to be allowable as being dependent on allowable amended claims 1 and 19, respectively.

All of the issues raised by the Examiner have been dealt with. In view of the foregoing, it is submitted that all the claims now pending in the application are allowable over the cited reference. An early Notice of Allowance is therefore respectfully requested.

Respectfully submitted,

Martin D. Moynihan
Registration No. 40,338

Date: February 7, 2006